Stoch Environ Res Risk Assess (2011) 25:353–361
DOI 10.1007/s00477-010-0373-4

ORIGINAL PAPER

Assessment of temporal and spatial variation of coastal water quality and source identification along Macau peninsula

Jinliang Huang · Manhim Ho · Pengfei Du

Published online: 9 February 2010

Abstract Cluster analysis (CA), discriminant analysis (DA) and principal component analysis (PCA) were comprehensively coupled to explore and identify the spatial and temporal variation and potential pollution sources in coastal water quality along Macau peninsula. The results show that the 12 months could be grouped into two periods, June–September and the remaining months, and the entire area divided into two clusters, one located at the western sides, and the other on the southeast and southern sides of the Macau peninsula. Through backward stepwise DA, pH, Cl^-, TSS, Color and TP, Chloride, Color, NH_4^+, DO, COD were discriminant variables of spatial and temporal variation, with 84.82 and 76.54% correct assignments, respectively. Fecal pollution, organic pollution and soil weathering are among the major sources for coastal water quality deterioration along Macau peninsula. This study illustrates that application of multivariate statistical techniques was beneficial to gain knowledge for further optimizing the monitoring network and controlling coastal water quality along Macau peninsula.

J. Huang (✉)
Joint Key Laboratory of Coastal Study, Xiamen University,
Xiamen 361005, China
e-mail: jlhuang@xmu.edu.cn

J. Huang
Environmental Science Research Center, Xiamen University,
Xiamen 361005, China

M. Ho · P. Du
Department of Environmental Science and Engineering,
Tsinghua University, Beijing 100084, China

M. Ho
The Official Provisional Municipal Council of Macau,
Macau, China

Keywords Cluster analysis · Discriminant analysis · Principal component analysis · Coastal water quality · Macau peninsula

1 Introduction

Coastal water quality has become one of the great environmental concerns from worldwide to regional scales, and is influenced by natural and anthropogenic disturbance, such as wastewater, runoff effluents, land reclamation, atmospheric deposition and climate change (Bowen and Depledge 2006; Kuppusamy and Giridhar 2006). Regular implementation of monitoring programs is recognized to be the essential step to characterize and control the coastal water pollution (Simeonov et al. 2003; Singh et al. 2004). However, many monitoring programs result in large and complicated data sets consisting of physical, chemical, biological and microbiological properties, which are difficult to analyze and interpret because of latent interrelationships among parameters and monitoring sites (Zhou et al. 2007a). In Macau, in order to evaluate the state of health of coastal water and its long-term change, the Drainage Services Department of Macau Civil Administration has set inshore monitoring sites along Macau peninsula since 2000 and produces large data sets with high complexity. It is therefore necessary to extract the meaningful information from this data set for effective coastal water quality management along Macau peninsula.

Multivariate techniques including cluster analysis (CA), discriminant analysis (DA) and principal component analysis (PCA) have been applied to a variety of environmental applications including assessment of spatiotemporal change of rainwater quality (Wunderlin et al. 2001;

Vazquez et al. 2003; Hamers et al. 2003; Ouyang et al. 2006), identification of rainfall runoff characteristics on different urban surfaces (Goonetillekea et al. 2005; Yamada et al. 1993; Huang et al. 2007), and evaluation of the spatiotemporal patterns of land-based pollution on coastal areas (Zhou et al. 2007a, b, c), etc. Obviously, previous studies provided valuable insights into the application of CA, DA and PCA techniques to environmental management and protection. However, comprehensive application of CA, DA and PCA to analysis of the coastal water quality along Macau peninsula regarding spatial–temporal variation and sources identification has not been conducted.

In this study, we analyze the coastal water quality from 22 inshore sampling sites during 2000–2005 by the integration of CA, DA and PCA methods. The objectives of this study are: (1) to extract latent information about the similarities or dissimilarities among the monitoring periods or sites; (2) and to identify pollution sources leading to spatiotemporal variations in water quality in Macau peninsula.

2 Materials and methods

2.1 Study area and monitoring sites

Located on the western side of the Pearl River estuary, Macau lies between latitudes 22°06′39″N and 22°13′06″N, and between longitudes 113°31′33″E and 113°35′43″E. Macau consists of Macau peninsula, Taipa island and Coloane island, covering 9.3, 6.7 and 13.2 km^2, respectively. Macau has a total population of 538,100, whereas nearly 95% of the population is located in the Macau peninsula (DSEC 2007).

The Drainage Services Officer of Macau civil administration has set inshore sampling sites along the Macau peninsula since 2000. Figure 1 shows the monitoring sites and sewer systems. The sampling sites are set near the overflow manholes of the sewer pipes. The inshore surface seawater was sampled at about 1 m underwater.

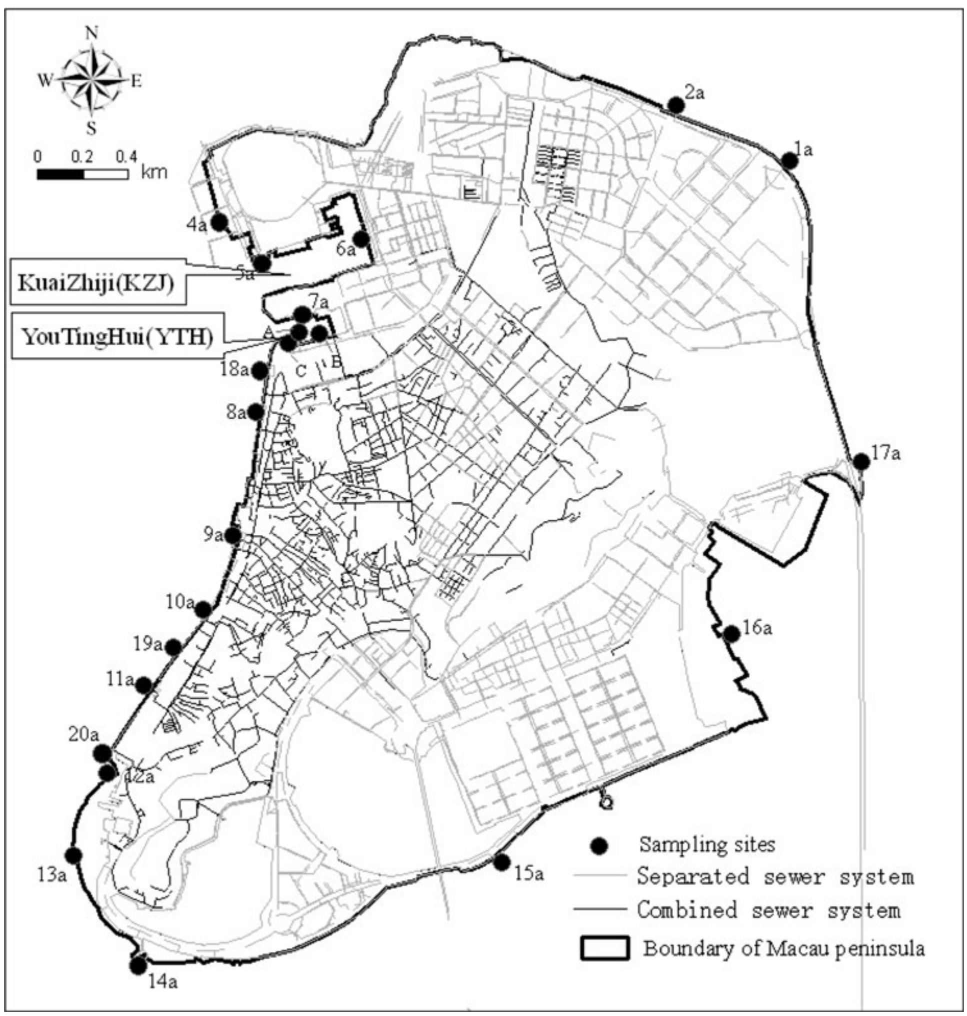

Fig. 1 Location of monitoring sites along Macau peninsula

2.2 Parameters and analytical methods

The data sets of coastal water quality from 22 monitoring sites consisted of 14 water-quality parameters monitored monthly over 6 years (Jan 2000–Dec 2005). The water quality parameters included Acoliform (*E.coli*), Fcoliform (*F.coli*), pH, color, turbidity, Electronic conductance (EC), Dissolved Oxygen (DO), chloride (Cl^-), Total Suspended Solids (TSS), nitrite nitrogen (NO_2^-), nitrate nitrogen (NO_3^-), ammonia nitrogen (NH_4^+), Total Phosphorus (TP), and Chemical Oxygen Demand (COD). Sampling and analysis for these parameters followed standard methods (APHA 1992). The 6-year data set consisted of 606 observations of coastal water quality in Macau peninsula.

2.3 Data treatment

Most multivariate statistical methods require variables to conform to the normal distribution, thus, the normality of the distribution of each variable was checked by analyzing kurtosis and skewness statistical tests before multivariable statistical analysis was conducted (Lattin et al. 2003). The original data demonstrated values of kurtosis ranging from 4.247 to 606.629 and skewness ranging from −0.714 to 24.626, indicating that distributions were far from normal with 95% confidence. Since most of the values of kurtosis or skewness were greater than 0, the original data were transformed in the form $x' = \log 10(x)$. After log-transformation, the kurtosis and skewness values ranges from −0.695 to 17.06 and −1.831 to 0.733, respectively. In the case of CA and PCA, all log-transformed variables were also *z*-scale standardized (the mean and variance were set to 0 and 1, respectively) to minimize the effects of difference units and variance of variables and to render the data dimensionless (Zhou et al. 2007c; Singh et al. 2004).

2.4 The multivariate techniques

In this study, CA, DA, and PCA were comprehensively coupled to carry out multivariate analysis for the data sets collected in 22 monitoring sites. A summary of principles of CA, DA, and PCA is described below.

2.4.1 Cluster analysis (CA)

Cluster analysis (CA) is an unsupervised pattern recognition method that divides a large group of cases into small groups or clusters of relatively similar cases that are dissimilar to other groups. Hierarchical CA, the most common approach, starts with each case in a separate cluster and joins the clusters together step by step until only one cluster remains (Lattin et al. 2003). The Euclidean distance usually gives the similarity between two samples, and a distance can be represented by the difference between transformed values of the samples (Otto 1998). In this study, hierarchical CA was performed on the standardized data using Ward's method with squared Euclidean distances as a measure of similarity. Ward's method uses analysis of variance (ANOVA) to calculate the distance between clusters to minimize the sum of squares of any two possible clusters at each step. Both temporal and spatial variations in water quality were determined from hierarchical CA using the linkage distance. The Ward method, based on hierarchical CA, is effective and popular (Wunderlin et al. 2001; Vazquez et al. 2003).

2.4.2 Discriminant analysis (DA)

Discriminant analysis (DA) is a method of analyzing dependence that is a special case of canonical correlation, and one of its objectives is to determine the significance of different variables, which can allow the separation of two or more naturally occurring groups. DA operates on original data, and the method constructs a discriminant function for each group as follows (Wunderlin et al. 2001; Zhou et al. 2007b):

$$f(G_i) = k_i + \sum_{j=1}^{n} w_{ij} p_{ij} \quad (1)$$

where i is the number of groups(G), k_i is the constant inherent to each group, n is the number of parameter numbers used to classify a set of data into a given group, and w_j is the weight coefficient, assigned by DA to a given parameter (p_j).

Backward stepwise DA was proved to be the most effective mode for reducing the dimensionality of the large data set (Wunderlin et al. 2001; Singh et al. 2004; Zhou et al. 2007b).

2.4.3 Principal component analysis (PCA)

Principal component analysis (PCA) is one of the most powerful and common techniques used for reducing the dimensionality of large sets of data without loss of information. PCA mathematically operates from the covariance matrix, which describes the dispersion of the multiple measured parameters, to obtain eigenvalues and eigenvectors. Linear combinations of the original variables and eigenvectors result in new variables, called principal components (PCs). Further rotation of the axis defined by PCA produces new groups of variable called varifactors (VFs). The basic features associated with PCA are data reduction and data grouping. Data reduction is obtained because we usually need only a few PCs to get a good description of the entire data set variability without loss of information. The use of PCA to water quality assessment has increased in recent years, mainly due to the need to obtain appreciable data reduction for analysis and decision-making (Helena et al. 2000).

The steps of the PCA include: (1) regulate the original data; (2) create a relative coefficient matrix; (3) calculate the eigenvalue and eigenvector; (4) calculate the contribution and the accumulative contribution ratio, and confirm the main component; (5) calculate the main component factors matrix of the pollution load.

3 Results and discussions

3.1 Temporal similarity and period grouping

An initial exploratory approach involved the use of hierarchical CA on standardized log-transformed data sets sorted by season. CA generated a dendrogram (Fig. 2), grouping the 12 months into two clusters at $(D_{link}/D_{max}) \times 100 < 23$, and the difference between the clusters was significant. Cluster 1 (the first period) included June, July, August, September, closely corresponding to the wet season (April–September). Cluster 2 (the second period) included the remaining months (January–May, October–December), approximately corresponding to the dry season in Macau. However, if the months had been empirically divided into spring (March–May), summer (June–September), autumn (October–December), and winter (January–February), or into dry/wet seasons, a mistake in grouping would have been made. In fact, Fig. 2 shows that the temporal patterns to water quality were not purely consistent with the four seasons or the dry/wet seasons.

83% of annual precipitation occurs during the period from April to September in Macau, so the grouping by CA basically corresponds to the dry/wet seasons. However, there will be little deviation error if the water quality is monitored only by the dry/wet season.

3.2 Spatial similarity and sites grouping

Considering the experience obtained from temporal CA, spatial CA was also used to identify similar monitoring sites. However, the influences of temporal differences on spatial-CA were considered. Both spatial similarity analysis for each temporal cluster and the integrated clusters (the first and second periods) were carried out, but the results were almost the same. Therefore, only the latter result is discussed. Spatial-CA produced a dendrogram, shown in Fig. 3, with second groups at $(D_{link}/D_{max}) \times 100 < 12$.

Group A consisted of a1, a2, a14, a15, a16, and a17, and Group B consisted of point a, point b, point c, a4–a13, and a18–a20. The group classifications varied with significance level, because the sites in these groups had similar features and natural backgrounds that were effected by similar sources. In Group A, six sites were located in the eastern and southern part of Macau peninsula. The water quality of

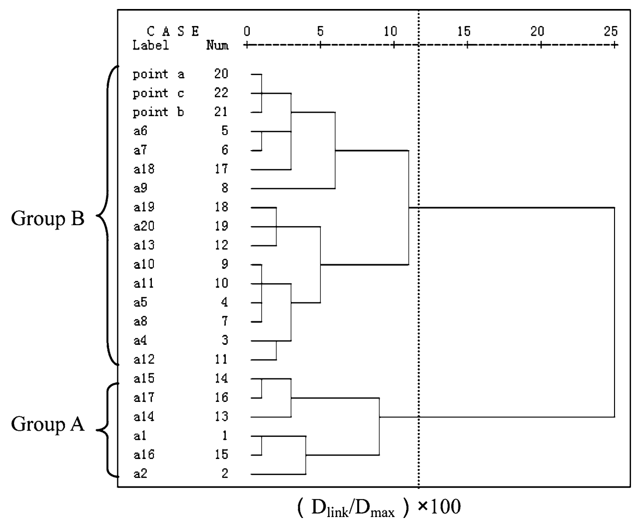

Fig. 3 Spatial cluster analysis of monitoring sites based on Ward's method

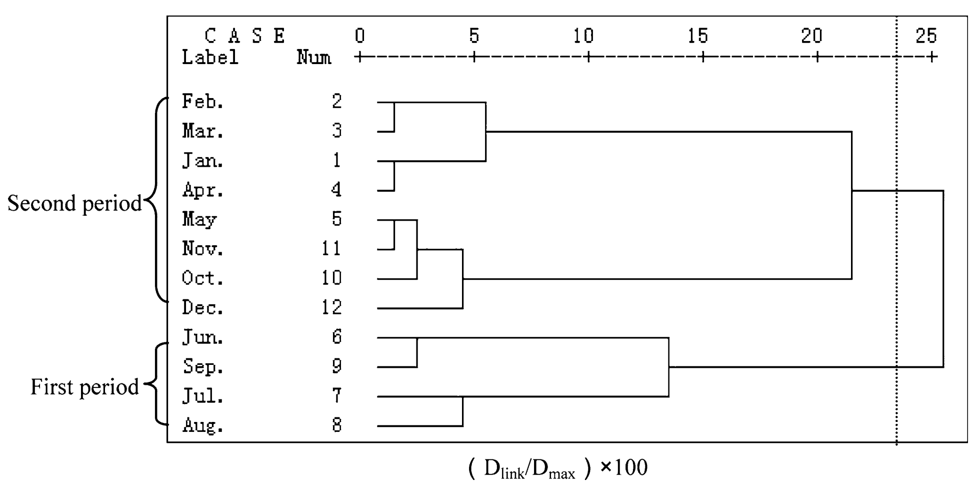

Fig. 2 Temporal cluster analysis of monitoring periods based on Ward's method

these sites is relatively good because these sites face the open ocean and the sewer system is separated (Fig. 1). In Group B, all sites were located in the western part of Macau peninsula. The water quality in Group B is poor because the region in the western side is the old urban area with a dense combined sewer system, where the major pollution sources were mainly from untreated domestic wastewater or combined sewer overflow (CSO). In contrast, coastal water quality in Group B is relatively poorer because this region (especially for YTH and KZJ shown in Fig. 1) does not face open coastal water, but a shadow watercourse located between Wanzai in Zhuhai district and Macau peninsula, which is not beneficial to the dilution of pollutants.

Hierarchical CA provided a useful classification of the coastal water in the Macau peninsula that can be used to design an optimal future spatial monitoring network with lower costs (Simeonov et al. 2003; Singh et al. 2004; Zhou et al. 2007b). According to the above results, the frequency of monitoring periods could only be selected from the first and second periods, and the number of monitoring sites could also be reduced and only chosen from Group A and Group B.

3.3 Temporal variation in water quality

Temporal variation in water quality parameters (Table 1) were evaluated using a period-parameter correlation matrix, which showed that most analyzed parameters were significantly correlated ($p < 0.05$) with period except *F.coli*, *E.coli*, pH, NO_3^-, and NO_2^-. Cl^- had the highest correlation coefficient ($R = 0.52$), followed by color ($R = 0.28$), DO ($R = 0.26$), EC ($R = -0.19$), TP ($R = 0.13$), NH_4^+ ($R = 0.12$), TSS ($R = -0.12$), Turbidity ($R = -0.11$), and COD ($R = -0.08$). These parameters accounted for the major temporal variation in water quality. The absence of a significant correlation of NO_3^-, NO_2^-, *F.coli*, *E.coli* with period indicates the contribution of anthropogenic pollution source in the coastal water along Macau peninsula.

Temporal variation in water quality was further evaluated using backward stepwise DA. Before running the temporal DA, the number of clusters needed to be decided, so the clusters based on temporal-CA were applied. The objectives of DA in this study were (1) to test the significance of discriminant functions and (2) to determine the most significant variables associated with the differences between the clusters.

Discriminant functions (DFs) and classification matrices (CMs) obtained from the backward stepwise modes of DA are shown in Tables 1 and 2. DA produced a CMs with close to 77% correct assignations using only six discriminant parameters. Thus, the temporal-DA results suggest that color, NH_4^+, Cl^-, TP, DO, and COD were the most significant parameters for discriminating between the first period and the second period and accounting for most of the expected temporal variation in water quality.

Box and whisker plots of the discriminant parameters are recognized by DA as being related to the temporal trend given in Fig. 4. The value of all discriminant parameter except color (including NH_4^+, Cl^-, TP, DO, and COD) is higher in the second period (January–May, October–December) than in the first period (June–September) (Fig. 4). The first period belongs to the wet season in Macau, when rainy weather enables stormwater runoff and soil loss occurs on many occasions, which makes the value of color relatively higher in the first period. Comparatively, there is less precipitation in the second period, coastal water is mainly influenced by the seawater, thus the value of Cl^- is higher in this period. Additionally, due to the land-based pollution effects discussed above, the coastal water quality in the second period has higher concentration of COD, TP, and NH_4^+.

3.4 Spatial similarity and sites grouping

The test significance in spatial-DA was calculated similarly to that for temporal-DA, as shown in Table 3. As shown in Table 3, the values of Wilks' lambda (0.61–0.63) was medium, which suggests that the spatial-DA in this study was relatively valid and effective.

Table 1 Classification matrix for backward DA of temporal variations

Number of parameters	Parameters	Wilks'	F	Percent correct
4	TP, Cl^-, color, NH_4^+	0.809	35.17800	63.04
5	TP, Cl^-, color, NH_4^+, DO	0.7977	30.20425	74.26
6	TP, Cl^-, color, NH_4^+, DO, COD	0.7905	26.28137	76.57

Table 2 Classification functions coefficients for backward stepwise DA of temporal variation

Parameters	First period coefficient[a]	First period coefficient[a]
Color	0.2932	0.1820
NH_4^+	0.4808	0.6052
Cl^-	0.0000	0.0003
TP	−1.3992	−0.4971
DO	0.8252	0.9833
COD	0.0085	0.0104
Constant	−4.5576	−5.1508

[a] Coefficients for different monitoring periods correspond to w_{ij} as defined in Eq. 1

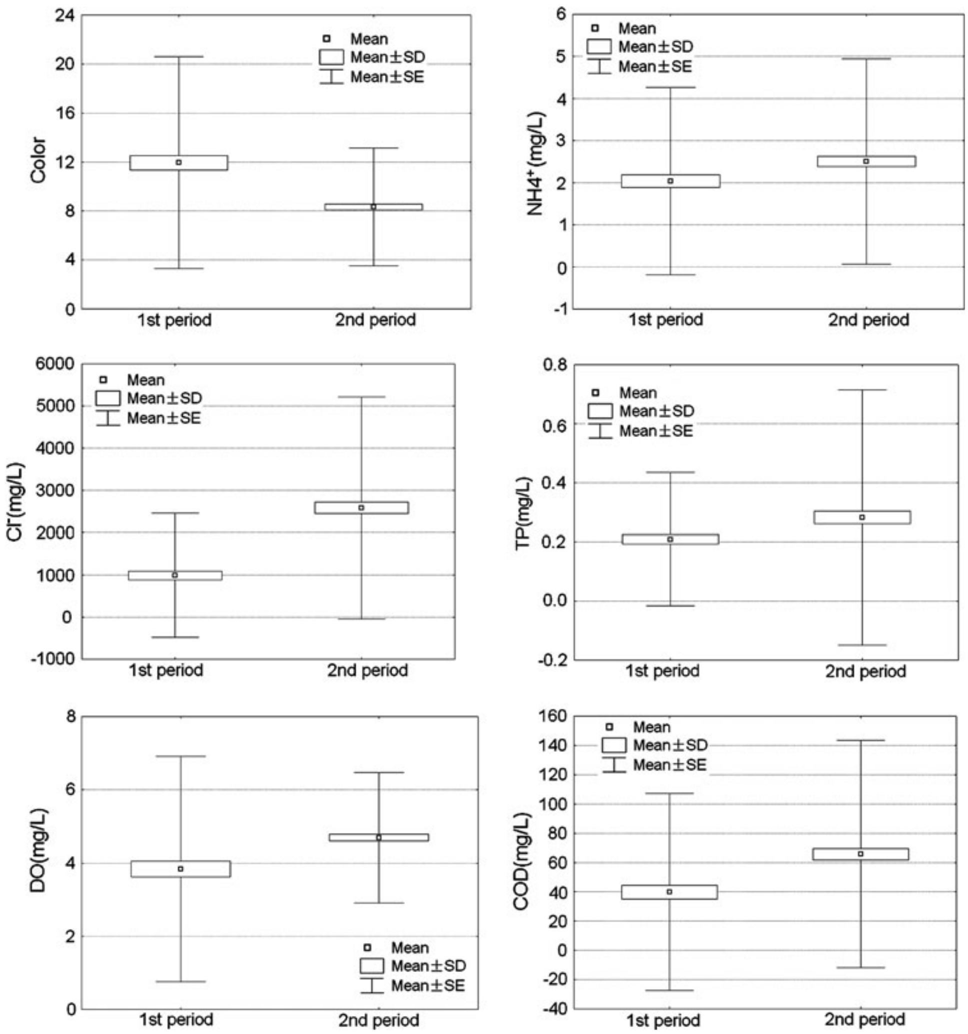

Fig. 4 Temporal variations of discriminant parameters derived from backward DA

Table 3 Classification matrix for backward DA of spatial variations

Number of parameters	Parameters	Wilks	*F*	Percent correct
3	pH, Cl^-, TSS	0.6285	117.5264	43.89
4	pH, Cl^-, TSS, color	0.6204	91.16763	86.47
5	pH, Cl^-, TSS, color, NH_4^+	0.6148	74.62154	84.82

Discriminant functions (DFs) and classification matrices (CMs) obtained from the backward stepwise modes of DA are shown in Tables 3 and 4. DA produced a CMs with close to 85% correct assignations using only five discriminant parameters. Thus, the spatial-DA results suggest that pH, Cl^-, TSS, color, and NH_4^+ were the most significant parameters for discriminating between Group A and Group B and accounting for most of the expected spatial variation in water quality.

Table 4 Classification functions coefficients for backward stepwise DA of spatial variation

Parameters	Group A coefficient[a]	Group B coefficient[a]
Color	0.3517	0.2881
pH	56.3645	54.8188
NH_4^+	1.6574	1.7871
Cl^-	0.0001	−0.0007
TSS	0.0229	0.0163
Constant	−221.822	−206.911

[a] Coefficients for different monitoring periods correspond to w_{ij} as defined in Eq. 1

Box and whisker plots of discriminating parameters identified by spatial backward stepwise DA were constructed to evaluate different patterns associated with spatial variation in water quality (Fig. 5). The average pH, color, Cl^- and TSS was higher in Group A than in Group B, whereas the average NH_4^+ was lower in Group A than in Group B.

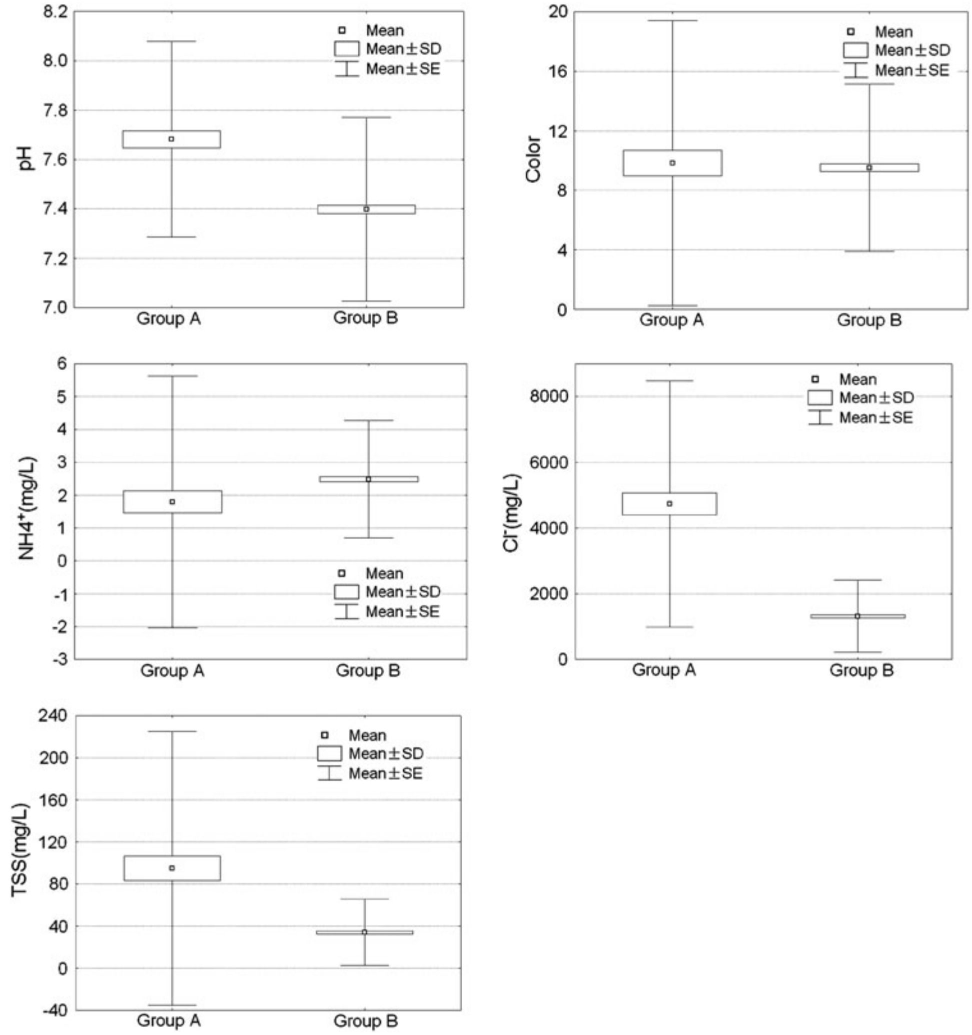

Fig. 5 Spatial variations of discriminant parameters derived from backward DA

As mentioned above, monitoring sites in Group A face the open coastal water, greatly suffering from seawater intrusion, which induces high levels of Cl^-. Meanwhile, wind and wave action in the estuary and the open coastal water conditions further leads to high levels of total suspended solid (TSS). In Group B, all sites are located in the western part of Macau peninsula (Fig. 1). The reason why coastal water quality in Group B is poorer than in Group A was discussed in Sect. 3.2. It is understandable that domestic wastewater from untreated or overflow sewer pipes constituted a major pollution source for the monitoring sites in Group B, which would cause a higher level of average NH_4^+ concentration and lower values of DO and pH. High levels of dissolved organic matter consume large amounts of oxygen, leading to anaerobic fermentation processes and the formation of ammonia and organic acids. Hydrolysis of these acidic materials causes a decrease of water pH value (Vega et al. 1998; Xi et al. 1999; Singh et al. 2004). Therefore, average pH value is lower in Group B than in Group A. The result of spatial-DA also supported the trends of discriminant parameters in water quality.

Based on the above results, backward stepwise DA proved to be a valuable tool to recognize the discriminant parameters in temporal and spatial variations of coastal water. Additionally, it was essential to strengthen the monitoring accuracy of pH, color, Cl^-, TSS, COD, DO, NH_4^+, and TP to clearly identify variations in future. Furthermore, pollution of Group B was relatively serious and should be controlled.

3.5 Identification of potential pollution sources in monitoring sites

PCA were further applied to standardized log-transformed data set (22 parameters) to examine differences between Groups A and Group B and identify the latent factors in different spatial variability, as shown in Table 5. The input data matrices (Variables × cases) for PCA were [14 × 126] for Group A and [14 × 480] for Group B. PCA of the two data sets evolved five PCs each for Group A and Group B, explaining 71.44 and 61.18% of the total variance

Table 5 Loadings of 14 measured variables on VARIMAX rotated factors of two spatial clusters by CA

Parameters	Group A					Group B				
	VF1	VF2	VF3	VF4	VF5	VF1	VF2	VF3	VF4	VF5
F.coli	**0.885**	0.168	0.034	0.209	−0.016	**0.858**	0.159	0.070	−0.079	−0.042
E.coli	**0.883**	0.168	0.078	0.311	−0.053	**0.865**	0.128	0.117	−0.101	−0.109
Color	0.393	0.256	−0.356	−0.087	−0.098	0.139	−0.246	−0.059	**0.693**	−0.060
Turbidity	0.113	**0.732**	−0.190	−0.078	0.082	−0.072	0.202	0.616	0.402	−0.050
pH	−0.147	−0.267	0.012	−0.046	0.021	−0.241	0.096	−0.002	−0.040	0.006
EC	−0.159	0.301	−0.011	−0.620	0.049	0.058	−0.276	**0.808**	−0.211	−0.002
NO_2^-	0.066	−0.038	0.217	0.033	**0.939**	−0.055	−0.141	−0.281	−0.209	**0.717**
NO_3^-	−0.257	0.112	−0.112	−0.322	0.543	−0.244	0.225	0.250	0.099	**0.782**
NH_4^+	0.415	0.108	0.112	**0.723**	−0.176	0.672	0.116	−0.340	0.129	−0.043
Cl^-	−0.118	−0.011	**0.795**	0.103	0.082	0.114	**0.792**	0.050	−0.072	0.121
TP	−0.014	0.636	0.092	0.681	0.047	0.554	0.158	−0.051	0.089	−0.175
DO	−0.187	−0.464	0.213	−0.027	0.016	−0.468	0.256	−0.104	−0.525	−0.084
COD	0.262	−0.077	**0.813**	0.017	0.017	0.124	**0.791**	−0.150	0.025	−0.081
TSS	−0.078	**0.766**	0.397	−0.227	0.013	−0.109	0.257	0.005	0.531	−0.068
Eigenvalue	4.16	2.40	2.16	1.43	1.24	2.48	1.57	1.24	1.09	0.99
% Total variance	26.10	15.08	13.53	8.98	7.76	21.59	13.67	10.79	9.52	8.60
Cumulative % variance	26.10	41.18	54.70	63.69	71.44	21.59	35.26	46.06	55.58	64.18

Significant factor loadings are bold faced

in respective water quality data sets. Corresponding VFs variables loading and variance explained are presented in Table 5.

As shown in Table 5, for Group A, among five VFs, VF1 explaining 26.10% of the total variance had strong positive loadings on *F.coli* and *E.coli*. Thus, VF1 represented domestic wastewater contaminated by fecal pollution (Jiang 2003). VF2 (15.08% of the total variance) had strong positive loadings on Turbidity and TSS. Thus, VF2 represents soil weathering and subsequent runoff (Singh et al. 2005), which is greatly related to the natural condition of estuary and open coastal water. VF3 (13.53% of the total variance) had strong positive loadings on Cl^- and COD. Thus VF3 represented organic pollution. VF4 (8.98% of total variance) and VF5 (7.76% of total variance) had strong positive loadings on NH_4^+ and NO_2^-, thus both represented nitrogenous nutrient pollution.

The pollution structure of Group B was similar to that of Group A. VF1, which explained 21.59% of the total variance, had strong positive loadings on *F.coli* and *E.coli*. Thus, VF1 represented domestic wastewater contaminated by fecal pollution (Jiang 2003). VF2 (13.67% of the total variance) had strong positive loadings on Cl^- and COD. Thus VF2 represented organic pollution and salt. VF3 (10.79% of the total variance) had strong positive loadings on electronic conductivity (EC). Thus, VF3 represented the natural influence of seawater). VF4 (9.52% of total variance) had medium positive loadings on color, indicating unidentified sources. VF5 (8.60% of total variance) had strong positive loadings on NO_3^- and NO_2^-, and thus represented nitrogenous nutrient pollution.

According to the results by PCA, domestic wastewater contaminated by fecal pollution was the most important potential pollution source both for Group A and Group B. Soil losses and subsequent runoff is also the major pollution source for Group A due to its natural location, namely, estuary and open coastal area, which easily suffered from the influence of wind and wave. Comparatively, organic pollution is another important latent pollution source for Group B, since domestic wastewater discharges from the dense combined sewer system always happens in the old urban areas around this group's sites. Such results were further validated by the spatial variation of coastal water quality by CA and DA in the above section.

4 Conclusions

Multivariate statistical analysis including CA, DA, and PCA was successfully applied to explore and identify temporal and spatial variation and potential pollution sources in coastal water quality along Macau peninsula, indicating that multivariate techniques are effective and useful for coastal water quality management.

Hierarchical CA grouped the 12 months into two periods, June–September and the remaining months, and the entire area divided into two clusters, one located at the western sides, and the other the southeast and southern sides of Macau peninsula. Through backward stepwise DA, pH, Cl^-, TSS, Color and TP, Cl^-, Color, NH_4^+, DO, COD were discriminant variables of spatial and temporal variation, with 84.82 and 76.54% correct assignments, respectively. Domestic wastewater contaminated by fecal pollution, organic pollution and soil losses are among the major sources for coastal water quality deterioration along Macau peninsula. Based on these findings, spatio-temporal pattern of in situ monitoring of coastal water quality along Macau peninsula was recognized. As a result, optimal future spatial–temporal monitoring network with lower cost should be designed in terms of date, sites, water quality parameters and potential land-based pollution sources when conducting in situ monitoring coastal water quality along Macau peninsula. This study illustrates that application of the multivariate statistical techniques was beneficial to gain knowledge for further optimizing the monitoring network and controlling coastal water quality along Macau peninsula.

Acknowledgements This research was funded by the Department of Science and Technology of Fujian province as a Talented Youth scholar in Fujian Province (No. 2007F3093). The authors also sincerely thank to Mrs. Lee Mue-heong (Official Provisional Municipal Council of Macau) and Dr. Zhou Feng (Environmental School, Peking University) for the suggestions and help of data analysis.

References

American Public Health Association, American Water Works Association, Water Environment Federation (1992) Standard methods for the examination of water and wastewater. American Public Health Association, Washington, DC

Bowen RE, Depledge MH (2006) Rapid assessment of marine pollution (RAMP). Mar Pollut Bull 53:638–646

Goonetillekea A, Thomasb E, Ginnc S, Gilbert D (2005) Understanding the role of land use in urban stormwater quality management. J Environ Manag 74:31–42

Hamers T, van den Brink PJ, Mos L, van der Linden SC, Legler J, Koeman JH, Murk AJ (2003) Estrogenic and esterase-inhibiting potency in rainwater in relation to pesticide concentrations, sampling season and location. Environ Pollut 123:47–63

Helena B, Pardo R, Vega M, Barrado E, Fernandez JM, Fernandez L (2000) Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis. Water Res 34:807–816

Huang J, Du P, Ao C, Lei M, Zhao D, Wang Z (2007) Multivariate analysis for stormwater quality characteristics identification from different urban surface types in Macau. Bull Environ Contam Toxicol 79:650–654

Jiang H (2003) Environmental water chemistry. Chemical Industry Press, Beijing

Kuppusamy MR, Giridhar VV (2006) Factor analysis of water quality characteristics including trace metal speciation in the coastal environmental system of Chennai Ennore. Environ Int 32:174–179

Lattin J, Carroll D, Green P (2003) Analyzing multivariate data. Duxbury Press, New York

Otto M (1998) Multivariate methods. In: Kellner R, Mermet JM, Otto M, Widmer HM (eds) Analytical chemistry. Wiley-VCH, Weinheim, Germany, 916 pp

Ouyang Y, Nkedi-kizza P, Wu QT, Shinde D, Huang CH (2006) Assessment of seasonal variations in surface water quality. Water Res 60:3800–3810

Simeonov V, Stratis V, Samara JA, Zachariadis G, Voutsa D, Anthemidis A (2003) Assessment of the surface water quality in Northern Greece. Water Res 37:4119–4124

Singh KP, Mali A, Mohan D, Sinha S (2004) Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India)—a case study. Water Res 38:3980–3992

Singh KP, Malik A, Sinha S (2005) Water quality assessment and apportionment of pollution sources of Gomti River (India) using multivariate statistical techniques—a case study. Anal Chim Acta 538:355–374

The Statistics and Census Service (DSEC) (2007) Yearbook of statistics 2007. Government Printing Bureau, Macao. http://www.dsec.gov.mo

Vazquez A, Costoya M, Pena RM, Garcia S, Herrero C (2003) A rainwater quality monitoring network: a preliminary study of the composition of rainwater in Galicia (NW Spain). Chemosphere 51:375–386

Vega M, Pardo R, Barrado E, Deban L (1998) Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. Water Res 32:3581–3592

Wunderlin DA, Diaz MDP, Ame MV, Pesce SF, Hued AC, Biston MLSA (2001) Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality. A case study: Suquia River Basin (Cordoba-Argentina). Water Res 35(12):2881–2894

Xi D, Sun Y, Liu X (1999) Environmental monitor. Higher Education Press, Beijing

Yamada K, Umehara T, Ichik A (1993) Study on statistical characteristics of nonpoint pollutants deposited in urban areas. Water Sci Technol 28:283–290

Zhou F, Guo H, Liu Y, Jiang Y (2007a) Chemometrics data analysis of marine water quality and source identification in Southern Hong Kong. Mar Pollut Bull 54:745–756

Zhou F, Huang GH, Guo H, Zhang W, Hao Z (2007b) Spatio-temporal patterns and source apportionment of coastal water pollution in eastern Hong Kong. Water Res 41:3429–3439

Zhou F, Liu Y, Guo H (2007c) Application of multivariate statistical methods to water quality assessment of the watercourses in Northwestern new territories, Hong Kong. Environ Monit Assess 132:1–13